Dispersion in time and space:
what propagating optical pulses in *time* (& not space) forces us to confront

P. Kinsler[*]

Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, United Kingdom.
(Dated: Friday 20th February, 2015)

I derive a temporally propagated uni-directional optical pulse equation valid in the few cycle limit. Temporal propagation is advantageous because it naturally preserves causality, unlike the competing spatially propagated models. The approach generates exact coupled bi-directional equations, which can be efficiently approximated down to a uni-directional form in cases where an optical pulse changes little over one optical cycle. It also also allows a direct term-to-term comparison of an exact bi-directional theory with an approximate uni-directional theory. Notably, temporal propagation handles dispersion in a different way, and this difference serves to highlight existing approximations inherent in spatially propagated treatments of dispersion. Accordingly, I emphasise the need for future work in clarifying the limitations of the dispersion conversion required by these types of approaches; since the only alternative in the few cycle limit may be to resort to the much more computationally intensive full Maxwell equation solvers.

NOTE: despite the implication of the claimed "substantial text overlap" flagged up by arXiv, the physical meaning of this work differs from that of arXiv:0810.5689 – it merely uses the same technique for decomposing waves into directional pieces, and as a consequence can have similar phrasing even when the physics is distinct.

I. INTRODUCTION

The importance of having simple and robust methods for the propagation of optical pulses has undergone increasing attention in recent years. This is due to the multitude of applications [1] in which increasing short pulses are relied upon to act as a kind of strobe-lamp to image ultrafast processes [2, 3], or even designer electric field profiles [4–6] to excite specific atomic or molecular responses. Alternatively, strong nonlinearity is used to construct equally wide-band but also temporally extended pulses – i.e. (white light) supercontinnua [7–9] – or even come full circle and use the strong nonlinearity to generate sub-structure that is again temporally confined, as in optical rogue waves [10]; or even the temporally and spatially localized filamentation processes [11–13].

It is clear, therefore, that progress towards diminishing pulse durations as well as their increasing spectral bandwidths, and higher pulse intensities are all either stretching existing pulse propagation models to their limits, or breaking them. In such regimes, we need to be sure that our numerical models still work, and have a clear idea of what has been neglected, and what the side-effects of those approximations are. In this paper I use a modern directional approach to show how a straightforward and relatively simple derivation allows a side-by-side comparison of exact and approximate propagation equations, whilst still providing the numerical and analytical convenience of a first-order wave equation.

However, unlike the typical approach in nonlinar optics, I consider propagation in time rather than along some chosen spatial axis [14, 15]. The resulting evolution equations for the wave profiles in space no longer give us access to the frequency spectrum [16], but instead we have the wavevec-tor spectrum, as well as a causally appropriate time evolution [17]. This requires an alternate approach to the temporal response of the propagation medium, which highlights the existing and somewhat poorly characterised approximations inherent in spatially propagated treatments of dispersion.

I start in section II by deriving a customized second order wave equation from Maxwell's equations, and reorganize it to define the material properties appropriately and set up the factorization stage. In section III, I use the method of factorization that allows us to construct an explicitly bi-directional model, and which is then reduced to the popular uni-directional limit in section IV. Section IV D discusses typical modifications that can be applied to the equations given in sections III and IV in order to and simplify them appropriately and compare them to existing models. Having derived this main result – the propagation equations and their various specialisations – I turn in V to the consequences of temporal propagation and the new perspective it provides on the handling of dispersive effects in this and standard (spatially propagated) models. This is then followed in section VI by a specific comparison between the ordinary spatially propagated nonlinear Schrödinger (NLS) equation – perhaps the most widely investigated equation in nonlinear optics – and its temporally propagated counterpart. Finally, in section VII, I present my conclusions.

II. SECOND ORDER WAVE EQUATION

Most optical pulse problems consider a uniform and source free dielectric medium. In such cases a good starting point is the second order wave equation for the electric field, which results from the substitution of the $\nabla \times \mathbf{H} = \partial_t \mathbf{D} + \mathbf{J}$ Maxwell's equation into $\nabla \times \mathbf{E} = -\partial_t \mathbf{B}$ (see e.g. [18]). Here, however, I want to follow the displacement field $\mathbf{D}$ (not $\mathbf{E}$) since it naturally occurs in conjuction with a time derivative – likewise,

[*]Electronic address: Dr.Paul.Kinsler@physics.org

the other important field is **B** rather than **H**. Thus we rexpress the curl Maxwell's equations as

$$\mu_0 \partial_t \mathbf{D} = \mu_0 \nabla \times \mathbf{H} - \mu_0 \mathbf{J} \quad = \nabla \times \mathbf{B} - \mu_0 \nabla \times \mathbf{M} - \mu_0 \mathbf{J}. \tag{2.1}$$

$$\varepsilon_0 \partial_t \mathbf{B} = -\varepsilon_0 \nabla \times \mathbf{E} - \varepsilon_0 \mathbf{K} \quad = -\nabla \times \mathbf{D} + \nabla \times \mathbf{P} - \varepsilon_0 \mathbf{K}. \tag{2.2}$$

Note that for both of these equations, we will also (as usual) want to specify the material response, whether dielectric or magnetic, in terms of a dielectric polarization $\mathbf{P} = \mathbf{D} - \varepsilon_0 \mathbf{E}$ and a magnetization $\mathbf{M} = \mathbf{B} - \mu_0 \mathbf{H}$. Either of these equations given above might be either in the spatial domain (with argument $\mathbf{r} = (r_x, r_y, r_z)$), or the wavevector domain (with argument $\mathbf{k} = (k_x, k_y, k_z)$).

Now we define the total magnetization $\mathbf{M}$ in terms of a component $\mathbf{M}_L$ linearly dependent on **B**, and another component $\mathbf{M}_B$ containing the rest. As a result, the excess polarization $\mathbf{M}_\mu \equiv \mathbf{M}_B$, which will typically include any nonlinearity, is also a function of **B** and not **H**. In traditional **H**-based spatially-propagated approaches, the temporal response ("dispersion") can be incorporated naturally, here it cannot. We have that

$$\mathbf{M}(\mathbf{r}) = \mathbf{M}_L(\mathbf{r}) + \mathbf{M}_B(\mathbf{r}) \quad = \alpha_\mu(\mathbf{r})\mathbf{B}(\mathbf{r}) + \mathbf{M}_B(\mathbf{r}), \tag{2.3}$$

which in the spatial frequency (wavevector **k**) domain is

$$\mathbf{M}(\mathbf{k}) = \mathbf{M}_L(\mathbf{k}) + \mathbf{M}_B(\mathbf{k}) \quad = \alpha_\mu(\mathbf{k}) \star \mathbf{B}(\mathbf{k}) + \mathbf{M}_B(\mathbf{k}), \tag{2.4}$$

since after a Fourier transform from **r** into **k**, the products become spatial convolutions ("$\star$"); which emphasises that anything wavevector dependent is necessarily a non-local concept. Note that this is the converse of the usual situation where a temporal response is described using time-domain convolutions, and these become products in the frequency domain.

Now we define the total polarization **P** in terms of a component $\mathbf{P}_L$ linearly dependent on **D**, and another component $\mathbf{P}_D$ containing the rest. As a result, the excess polarization $\mathbf{P}_\varepsilon \equiv \mathbf{P}_D$, which will typically include any nonlinearity, is also a function of **D**. We have that

$$\mathbf{P}(\mathbf{r}) = \mathbf{P}_L(\mathbf{r}) + \mathbf{P}_D(\mathbf{r}) \quad = \alpha_\varepsilon(\mathbf{r})\mathbf{D}(\mathbf{r}) + \mathbf{P}_D(\mathbf{r}), \tag{2.5}$$

which in the spatial frequency (wavevector **k**) domain is

$$\mathbf{P}(\mathbf{k}) = \alpha_\varepsilon(\mathbf{k}) \star \mathbf{D}(\mathbf{k}) + \mathbf{P}_D(\mathbf{k}). \tag{2.6}$$

With these definitions for $\partial_t \mathbf{D}$, $\partial_t \mathbf{B}$, and for **M** and **P** set up, we can now proceed with the derivation of a second order wave equation for **D** –

$$\mu_0 \partial_t \mathbf{D} = \nabla \times \mathbf{B} - \mu_0 \nabla \times \mathbf{M}_L - \mu_0 \nabla \times \mathbf{M}_B - \mu_0 \mathbf{J} \tag{2.7}$$

$$= \nabla \times \mathbf{B} - \nabla \times \alpha_\mu \mathbf{B} - \mu_0 \nabla \times \mathbf{M}_B - \mu_0 \mathbf{J} \tag{2.8}$$

$$= \nabla \times \mathbf{B} - \alpha_\mu \nabla \times \mathbf{B} - \left(\nabla \alpha_\mu\right) \times \mathbf{B} - \mu_0 \nabla \times \mathbf{M}_B - \mu_0 \mathbf{J} \tag{2.9}$$

$$= \nabla \times \left[1 - \alpha_\mu\right] \mathbf{B} - \mu_0 \nabla \times \mathbf{M}_B - \mu_0 \mathbf{J}. \tag{2.10}$$

Now take the time derivative of this equation, use the fact that α_μ is independent of time, define $a_\mu = 1 - \alpha_\mu$ and $a_\varepsilon = 1 - \alpha_\varepsilon$; and then substitute for $\partial_t \mathbf{B}$, so that

$$\varepsilon_0 \mu_0 \partial_t^2 \mathbf{D} = \varepsilon_0 \nabla \times a_\mu \partial_t \mathbf{B} - \varepsilon_0 \mu_0 \nabla \times \partial_t \mathbf{M}_B - \varepsilon_0 \mu_0 \partial_t \mathbf{J} \tag{2.11}$$

$$c^{-2} \partial_t^2 \mathbf{D} = -\nabla \times a_\mu \left[\nabla \times \mathbf{D} - \nabla \times \mathbf{P} + \varepsilon_0 \mathbf{K}\right] - c^{-2} \nabla \times \partial_t \mathbf{M}_B - c^{-2} \partial_t \mathbf{J} \tag{2.12}$$

$$\partial_t^2 \mathbf{D} = -c^2 \nabla \times a_\mu \left[\nabla \times \mathbf{D} - \nabla \times \alpha_\varepsilon \mathbf{D} - \nabla \times \mathbf{P}_D + \varepsilon_0 \mathbf{K}\right] - \nabla \times \partial_t \mathbf{M}_B - \partial_t \mathbf{J} \tag{2.13}$$

$$= -c^2 \nabla \times a_\mu \left[\nabla \times \left(1 - \alpha_\varepsilon\right) \mathbf{D} - \nabla \times \mathbf{P}_D + \varepsilon_0 \mathbf{K}\right] - \nabla \times \partial_t \mathbf{M}_B - \partial_t \mathbf{J} \tag{2.14}$$

$$= -c^2 \nabla \times a_\mu \nabla \times a_\varepsilon \mathbf{D} + c^2 \nabla \times a_\mu \nabla \times \mathbf{P}_D - c^2 \nabla \times a_\mu \varepsilon_0 \mathbf{K} - \nabla \times \partial_t \mathbf{M}_B - \partial_t \mathbf{J} \tag{2.15}$$

Remember that in this picture, all field properties are specified as functions of our chosen primary field **D**; notably, the magnetic field **B** which is usually the main argument for $\mathbf{M}_B$ must be determined from the calculated **D** using the relevant Maxwell's equation (i.e. eqn. (2.2)).

A. Simple case: isotropic & homogeneous

Since trying to treat all the details of eqn. (2.15) correctly leads to excessively complicated expressions, I will first simplify it down into the case where (a) the material's reference properties a_μ, a_ε do not vary in space, and (b) the effective monopole current is zero ($\mathbf{K} = 0$). Since (magnetic) monopoles do not exist, and all non reference properties can still be encoded in $\mathbf{P}_D$, $\mathbf{M}_B$, or **J**, this simplification does not (of itself) impose any approximation. However, the first simplification means that our reference properties cannot incorporate any exact knowledge about the spatial structure of the propagation medium[1]. This restriction on obtaining the best possible match between the reference behaviour and the exact behaviour of the propagation medium will typically affect later approximations, where we assume deviations from the reference behaviour are small. One additional advantage of selecting constant reference parameters α_μ and α_ε is that we can replace them with matrices to allow for e.g. birefringence and cross-polarization couplings.

The simplified second order wave equation for changes in the displacement field $\mathbf{D}(\mathbf{r})$ as it propagates forward in time

---

[1] This differs from the situation present in the complementary spatially propagated approach [14], where the medium's linear temporal response typically *can* be subsumed into the reference behaviour. Nevertheless, in neither case can the *spatial* properties be incorporated exactly into the reference behaviour; although approximations – notably that of a so-called "spatial dispersion" – can be made.

is

$$\partial_t^2 \mathbf{D} = -c^2 a_\mu a_\varepsilon \nabla \times \nabla \times \mathbf{D} \quad + c^2 a_\mu \nabla \times \nabla \times \mathbf{P}_D$$
$$- \nabla \times \partial_t \mathbf{M}_B - \partial_t \mathbf{J} \qquad (2.1)$$
$$= +c^2 a_\mu a_\varepsilon \nabla^2 \mathbf{D} - c^2 a_\mu a_\varepsilon \nabla \nabla \cdot \mathbf{D}$$
$$- c^2 a_\mu \nabla^2 \mathbf{P}_D + c^2 a_\mu \nabla \nabla \cdot \mathbf{P}_D$$
$$- \nabla \times \partial_t \mathbf{M}_B - \partial_t \mathbf{J}. \qquad (2.2)$$

Note that the retention of the $\nabla \cdot \mathbf{D}$ and $\nabla \cdot \mathbf{P}_D$ terms allows us to include charge effects, as is needed when discussing high-power nonlinear situations such as filamentation [19].

In the spatial Fourier "wavevector" regime, we replace $\nabla \to \imath \mathbf{k}$, so that $k^2 = \mathbf{k} \cdot \mathbf{k}$, and reorganize slightly. The equation for changes in the displacement field $\mathbf{D}(\mathbf{k})$ as it propagates forward in time is

$$\partial_t^2 \mathbf{D} - c^2 a_\mu a_\varepsilon k^2 \mathbf{D} = -c^2 a_\mu k^2 \mathbf{P}_D - \partial_t \mathbf{J} - \imath \mathbf{k} \times \partial_t \mathbf{M}_B$$
$$- c^2 a_\mu a_\varepsilon \mathbf{k} \mathbf{k} \cdot \mathbf{D} + c^2 a_\mu \mathbf{k} \mathbf{k} \cdot \mathbf{P}_D \quad (2.3)$$
$$\partial_t^2 \mathbf{D} - \Omega^2 \mathbf{D} = +\Omega^2 a_\varepsilon^{-1} \mathbf{P}_D - \partial_t \mathbf{J} - \imath \mathbf{k} \times \partial_t \mathbf{M}_B$$
$$- \Omega^2 k^{-2} \mathbf{k} \mathbf{k} \cdot \mathbf{D} + \Omega^2 k^{-2} a_\varepsilon^{-1} \mathbf{k} \mathbf{k} \cdot \mathbf{P}_D,$$
$$(2.4)$$

where we have defined a reference frequency Ω for the propagation, based on the reference material properties, i.e.

$$\Omega^2(\mathbf{k}) = c^2 a_\mu a_\varepsilon k^2. \qquad (2.5)$$

We could – if we believed we understood the spatial properties of the propagation medium sufficiently well – simply alter this definition by incorporating an alternate $\mathbf{k}$ dependence that mimics an assumed spatial dispersion. However, it needs to be understood that such a step is distinctly *ad hoc*, since it lacks the convolutions over $\mathbf{k}$ necessary for any accurate representation of the material structure. But, that said, such an approach may still have considerable practical utility; and indeed is functionally equivalent to how other methods handle spatial dispersion, as a non-local (and hence, strictly speaking, a non-causal) process: e.g. those that import a waveguide dispersion calculated from the (spatial) modes given by the guide's transverse cross-section. To allow for this possibility, in the following I will replace the physically rigorous magnitude k argument for Ω with $\mathbf{k}$, its vector counterpart[2]. This allows for not only more complicated assumed spatial dispersions, but also additional orientation dependence. However, these benefits (only) arise as a result of approximation. Strictly speaking, both on the grounds covered above, and ones taking stricter view of causality [16], any notion of spatial dispersion is incompatible with locality.

Unlike for the spatially propagated scheme in [14], we have not included temporal response of the propagation medium in

---

[2] In some cases, depending on the approximate model chosen, this new $\Omega(\mathbf{k})$ definition may also need to be convolved with other parts of the expression, e.g. $\Omega^2(k)\mathbf{P}_D(\mathbf{k}) \to \Omega^2(\mathbf{k}) \star \mathbf{P}_D(\mathbf{k})$, although such an additional complication is likely to prevent us obtaining much conceptual or mathematical benefit. Consequently, in what follows I do not allow for such convolutions.

our definition of the directional fields, since it is incompatible with our temporal propagation. To include this we must model it explicitly as is done in Maxwell solvers such as FDTD [20, 21]; or approximate it as if it were a spatial effect (see e.g. [22], or the discussions in the later section V.

III. FACTORIZATION

I now factorize the wave equation [14, 23], a process which neatly avoids some of the approximations necessary in traditional approaches. Factorization takes its name from the fact that the LHS of eqn. (2.2) or (2.4) is a simple difference of squares which might be factorized, indeed this is what was done in a somewhat ad hoc fashion by Blow and Wood in 1989 [24]. Since the factors are just of the form $(\partial_t \mp \imath k)$, each by itself looks like a forward (or backward) directed wave equation. It therefore allows us to define a pair of counter-propagating Greens functions, and so divide the second order wave equation into a bi-directional pair of coupled first order wave equations. In addition, it allows us to directly compare the exact bi-directional and approximate uni-directional theories term for term, whereas in other approaches the backward parts simply vanish and are not directly available for comparison. A short summary of the factorization method is given in appendix A.

To re-iterate an important point – the choice of a_ε, a_μ and therefore $\Omega(\mathbf{k})$ in eqn. (2.4) defines the specific Greens functions used; it therefore also defines the basis upon which we will then propagate the displacement field $\mathbf{D}$.

A. Bi-directional wave equations

A pair of bi-directional wave equations suggests similarly bi-directional fields, so I split the electric displacement field into forward ($\mathbf{D}^+$) and backward ($\mathbf{D}^-$) directed parts, with $\mathbf{D} = \mathbf{D}^+ + \mathbf{D}^-$. In the following equations, remember that $\mathbf{P}_D$, $\mathbf{M}_B$, etc all depend on the full field $\mathbf{D}$, and *not* only one partial field (e.g. only $\mathbf{D}^+$). This is an important point since we see they drive both forward and backward evolution equations equally.

The coupled bi-directional first order wave equations for the directed fields $\mathbf{D}^\pm$ are *propagated* forward in time whilst

being *evolved* forward and backward in space are

$$\partial_t \mathbf{D}^\pm(\mathbf{k}) = \pm\iota\Omega(\mathbf{k})\mathbf{D}^\pm(\mathbf{k}) \quad \pm \frac{\iota\Omega^2(\mathbf{k})a_\varepsilon^{-1}}{2\Omega(\mathbf{k})}\mathbf{P}_D(\mathbf{D},t;\mathbf{k})$$

$$\mp \frac{\dot{\mathbf{J}}(t;\mathbf{k})}{2\Omega(\mathbf{k})} \mp \frac{\iota}{2\Omega(\mathbf{k})}\mathbf{k}\times\dot{\mathbf{M}}_B(\mathbf{D},t;\mathbf{k})$$

$$\mp \frac{\iota\Omega^2(\mathbf{k})}{2\Omega(\mathbf{k})k^2}\mathbf{k}\mathbf{k}\cdot\mathbf{D}(\mathbf{k}) \pm \frac{\iota\Omega^2(\mathbf{k})a_\varepsilon^{-1}}{2\Omega(\mathbf{k})k^2}\mathbf{k}\mathbf{k}\cdot\mathbf{P}_D(\mathbf{D},t;\mathbf{k})$$

$$(3.1)$$

$$\partial_t \mathbf{D}^\pm(\mathbf{k}) = \pm\iota\Omega(\mathbf{k})\mathbf{D}^\pm(\mathbf{k}) \quad \pm \frac{\iota\Omega(\mathbf{k})}{2a_\varepsilon}\mathbf{P}_D(\mathbf{D},t;\mathbf{k})$$

$$\mp \frac{\dot{\mathbf{J}}(t;\mathbf{k})}{2\Omega(\mathbf{k})} \mp \frac{\iota}{2\Omega(\mathbf{k})}\mathbf{k}\times\dot{\mathbf{M}}_B(\mathbf{D},t;\mathbf{k})$$

$$\mp \frac{\iota\Omega(\mathbf{k})}{2}\frac{\mathbf{k}\mathbf{k}}{k^2}\cdot\mathbf{D}(\mathbf{k}) \pm \frac{\iota\Omega(\mathbf{k})}{2a_\varepsilon}\frac{\mathbf{k}\mathbf{k}}{k^2}\cdot\mathbf{P}_D(\mathbf{D},t;\mathbf{k})$$

$$(3.2)$$

Here I have replaced the partial time derivatives on the RHS with over-dots, this being to emphasise that if our expression is to remain strictly causal [17], then our models for current $\mathbf{J}$ and magnetization $\mathbf{M}_B$ must return their *time-derivatives* as explicit functions of known quantities (here, typically $\mathbf{D}$ or t).

Note that factorized equations can be rebuilt into a single second-order equation by taking the sum and difference, then substituting one into another with the assistance of a derivative w.r.t. t (see section IV.B of [23]).

A remaining clarification required is to understand the meaning of "forward" and "backward" in the context of three dimensional space – this being less straightforward than in the spatially propagated case, where one spatial axis is selected[3], and forward and backward refers to time t. Here, we should understand "forward" and "backward" as being with respect to a particular choice of $\mathbf{k}$, and those individual vector components, that we are interested in.

B. Propagation, evolution, and directed fields

Note that since our solutions of the wave equations enforce *propagation* towards later times t, the fields $D^\pm(t)$ are *directed* forwards and backwards in space; these fields then *evolve* forwards and/or backwards in space as t increases. Note that I use this terminology (propagated, directed, evolved) throughout this paper to mean these three specific things.

When examining the wave equation eqn. (3.2) which evolves the directed fields $\mathbf{D}^\pm$ as they propagate forward in t, we see that the RHS has two types of terms: which I label the "reference" and "residual" parts [25].

Reference evolution (or *"underlying evolution"*) is that given by $\pm\iota\Omega(\mathbf{k})\mathbf{D}^\pm$ term, and is determined by our chosen a_ε and a_μ, and any additional ad hoc refinements. By itself, it would describe an ordinary oscilliatory evolution where the

field oscillations in $\mathbf{D}^\pm(\mathbf{r})$ would move forward (+) or backwards (−) in space. This is analogous to the choice of reference when constructing directional fields [26], but here is done for temporal rather than spatial propagation [27].

Residual evolution accounts for the discrepancy between the true evolution and the underlying/reference evolution, and is every part of the material response *not* included in a_ε, a_μ (and hence $\Omega(\mathbf{k})$). i.e. it is all the remaining terms on the RHS of eqn. (3.2). These typically include any non-linear polarization, or spatially or orientationally dependent linear terms; they are analogous to the correction terms used in directional fields models. In the language used by Ferrando et al. [23], these residuals are "source" terms. Although we might hope they will be only a weak perturbation, so that we might make the (desirable) uni-directional approximation discussed later, the factorization procedure is nevertheless valid for *any* strength.

C. Reference evolution: choice of Ω and the resulting $D^\pm$

I now examine how the choice of Ω affects the relative sizes of the forward and backward directed D^+ and D^-. To do this I consider the simple example of a medium for which the field is known to propagate with frequency ω; but for demonstration purposes we choose a reference evolution determined by a frequency Ω that is different from ω. E.g., for a linear isotropic medium we could exactly define $\omega^2 = \Omega^2 + \Delta^2$; but in general we would just have some residual (source) term $\mathcal{Q}$. This means that our definitions of forward and backward *directed* fields do not exactly correspond to what the wave equation will actually *evolve* forward and backward as we *propagate* towards larger t.

The second order wave equation is $(\partial_t^2 + \Omega^2)\mathbf{D} = -\mathcal{Q}$, which in the linear case has $\mathcal{Q} = \Delta^2 D$, so that $(\partial_t^2 + \omega^2)D = 0$. The factorization in terms of Ω is then

$$\partial_t \mathbf{D}^\pm = \pm\iota\Omega\mathbf{D}^\pm \pm \frac{\iota\mathcal{Q}}{2\Omega}. \qquad (3.1)$$

Now if we select the case where our field D only evolves forward, we know that $\mathbf{D} = \mathbf{D}_0\exp[\iota\omega t]$. Consequently $D^\pm$ must have matching oscillations: i.e. $\mathbf{D}^\pm = \mathbf{D}_0^\pm\exp[\iota\omega t]$, even though D^- is directed backwards. Substituting these into eqn. (3.1) gives

$$\mathbf{D}_0^- = \frac{\Omega - \omega}{\Omega + \omega}\mathbf{D}_0^+, \qquad (3.2)$$

which specifies how much $\mathbf{D}^-$ we need to combine with $\mathbf{D}^+$ so that our pulse evolves forward; since the $\mathbf{D}^-$ will be dragged forward by its coupling to $\mathbf{D}^+$. This interdependent $\mathbf{D}^\pm$ behaviour is generic – no matter what the origin of the discrepancy between Ω and the true evolution of the field (i.e. the residual/source terms such as mismatched linear behaviour, nonlinearity, etc): some non-zero backward directed field $\mathbf{D}^-$ must exist but still evolve forwards with $\mathbf{D}^+$. Analogous behaviour can be seen in the directional fields approach of Kinsler et al. [26].

---

[3] Typically a cartesian axis, although other choices are possible [15].

Usually we hope that this residual $\mathbf{D}^-$ contribution is small enough so that it can be neglected. If we assume $\mathbf{D}^- \simeq 0$, then we find that $\omega \simeq \Omega + \Delta^2/2\Omega$, which is just the expansion of $\omega = (\Omega^2 + \Delta^2)^{1/2}$ to first order in Δ^2/Ω^2. Following this, we find that eqn. (3.2) then says that $\mathbf{D}_0^- \simeq (\Delta^2/4\Omega^2)\mathbf{D}_0^+$, which has come full circle and provided us with the scale on which $\mathbf{D}^-$ can be considered negligible. Outside the restricted (linear) case where we know Δ^2, the true frequency ω might be difficult to determine, and in nonlinear propagation any local estimate of the frequency can change as the pulse propagates.

There is a further important point to notice: if we choose $\Omega = \Omega(k)$ with a wave-vector dependence, then we see that the source-like terms inherit that dispersion; and this is almost inevitable, since typically $\Omega \sim ck$. This means, for example, that even if we started with nonlinear polarization model without any k dependence, our factorized equations will have a k dependence on the nonlinear terms. This is the complementary process to what happens in the spatially propagated directional fields approach of Kinsler et al. [26], where choosing a dispersive reference has an equivalent effect on the correction terms.

IV. UNI-DIRECTIONAL WAVE EQUATIONS

Making only a single, well defined class of approximation I can now reduce the exact coupled bi-directional evolution of $\mathbf{D}$ down to a single uni-directional first order wave equation. As with the complementary spatially propagated theory [14], I do not require a moving frame, a smooth envelope, or to assume inconvenient second order derivatives are somehow negligible, as was necessary in traditional treatments. The approximation is that the residual terms are weak compared to the (reference) $\pm \imath \Omega \mathbf{D}$ term – e.g. weak nonlinearity or orientation dependence. This enables me to assert that if I start with $\mathbf{D}^- = 0$, then $\mathbf{D}^-$ will remain negligible. In this context, "weak" means that no significant change in the backward field is generated in a distance shorter than one wavelength ("slow evolution"); and that small effects do not build up gradually over propagation times of many periods ("no accumulation").

Slow evolution is where the size of the residual terms is much smaller than that of the underlying linear evolution – i.e. smaller than $\Omega \mathbf{D}$. This allows us to write down straightforward inequalities which need to be satisfied. It is important to note the close relationship between these and a good choice of Ω, as discussed in subsection III C. If Ω is not a good enough match, there always be significant contributions from both forward and backward directed fields; and even if nothing ends up *evolving* backwards, an ignored backward directed field will result in (e.g.) miscalculated nonlinear effects, since the total field $\mathbf{D} = \mathbf{D}^+ + \mathbf{D}^-$ will be wrong.

No accumulation occurs when the evolution of any backward directed field $\mathbf{D}^-$ is dominated by its coupling via the residual terms to the forward directed field $\mathbf{D}^+$; and not by its preferred underlying backward evolution. No accumulation means that forward evolving field components do not couple to field components that evolve backwards; this the typical behaviour since the phase mismatch between forward evolv-

ing and backward evolving components is $\sim 2\Omega$; in essence it is comparable to the common rotating wave approximation (RWA). This rapid relative oscillation means that backward evolving components never accumulate, as each new addition will be out of phase with the previous one; it is not quite a "no reflection" approximation, but one that asserts that the many micro-reflections do not combine to produce something significant. An estimate of the terms required to break this approximation are given in appendix B; generally speaking this is a much more robust approximation than the slow evolution one. Of course, a periodic temporal modulation of the medium will give periodic residual terms, and these can be engineered to force phase matching. In spatially propagated analyses, where we talk of spatial (wavevector) phase matching and not this temporal (frequency) phase matching [28], this would be a periodicity based on a relatively small phase mismatch (see e.g. quasi phase matching in Boyd [29]); but might even go as far as matching the backward wave (see e.g. [30]).

It is also important to note that the same small size of perturbation from the residual terms *can* accumulate on the forward evolving field components (or, indeed, the backward perturbation on the backward evolving field components). Although the magnitude of the residual terms acting on the forward and backward field evolution are identical, forward evolving components of the residuals can accumulate on the forward evolving field because they are phase matched; whereas backward residuals are not, and rapidly average to zero.

A. Residual terms

The handling and criteria for different types of residual terms has been discussed already [14] for the spatially propagated case. While it would be straightforward to repeat that level of detail here, in fact the basic process is virtually identical except for the different roles played by the t and $\mathbf{r}$ arguments. Accordingly I will consider only the most commonly considered residuals – the dielectric scalar and vector terms as well as currents. Those interested in magnetic or magneto-electric effects can adapt the process to their own systems in a similar manner.

The relevant total polarization $\mathbf{P}_D$ can then be decomposed into pieces before seeing how each might satisfy the slow evolution criteria. Thus we write

$$\begin{aligned} \mathbf{P}_D(\mathbf{D}, \mathbf{r}, t) &= a_\varepsilon(\mathbf{D}, \mathbf{r}, t)\mathbf{D}(\mathbf{r}, t) + \mathbf{U}_D(\mathbf{D}, \mathbf{r}, t) \\ &= a_L(\mathbf{r}, t) * \mathbf{D}(\mathbf{r}, t) + a_N(\mathbf{D}, \mathbf{r}, t) * \mathbf{D}(\mathbf{r}, t) \\ &\quad + \mathbf{U}_L(\mathbf{r}, t) * \mathbf{D}(\mathbf{r}, t) + \mathbf{U}_N(\mathbf{D}, \mathbf{r}, t). \quad (4.1) \end{aligned}$$

The part which is scalar in nature is represented by a_ε, and might contain (non-reference) linear parts and time response (a_L), perhaps a convolution ("$*$") over the past, or angle dependence; it might also contain nonlinear contributions a_N such as the third order Kerr nonlinearity with $a_N \mathbf{D} \propto (\mathbf{D} \cdot \mathbf{D})\mathbf{D}$. The vector part $\mathbf{U}_D$, if present, could be from e.g. a second order nonlinearity, which couples the ordinary and extraordinary field polarizations. Note that this description of the

material parameters does not restrict them in any way, and can e.g. include any order of nonlinearity.

B. Slow evolution?

Now I will treat each possible residual term in order, where the oppositely directed field is negligible: i.e., for $\mathbf{D}^\pm$, we have that $\mathbf{D}^\mp \simeq 0$. where the scalar ε_L contains the linear response of the material that is both isotropic and lossless (or gain-less); since here it is a time-response function, it is convolved with the electric displacement field $\mathbf{D}$. Note that the field vector $\mathbf{D}$, and indeed the material parameter ε_L are all functions of time t and space $\mathbf{r} = (x, y, z)$; the polarization $\mathbf{P}_\varepsilon$ and its components ϕ_ε, $\mathbf{V}_\varepsilon$ are a functions of time t, space $\mathbf{r}$, and the field $\mathbf{D}$.

Below I will refer to field components D_i and $D_i^\pm$, where $\mathbf{D}^+ + \mathbf{D}^- = \mathbf{D} \equiv (D_x, D_y, D_z)$ and $i \in \{x, y, z\}$; also to wave vector components k_i from $\mathbf{k} = (k_x, k_y, k_z)$, with $k_\perp^2 = k_x^2 + k_y^2$. However, note that in the constraints below, that k_j is also used as a substitute symbol to represent any one of k_x, k_y, or k_z.

Any corrections due to imperfect choice of Ω will appear in the $\mathbf{P}_D$ polarization correction term. However, unlike in the spatially propagated case, there is no diffraction-like term – here, that is quite naturally part of the reference behaviour.

Firstly, scalar polarization terms ϕ, which can be either linear ($\phi_L(k)$) or nonlinear ($\phi_N(\mathbf{D}, k)$). These might encode e.g. the time-response of the medium (and hence its dispersion), or be some nonlinearity response such as the third-order Kerr nonlinearity. Since these are locally correlated in space, in the wavevector picture they include convolutions over k, and so give us the criterion

$$\frac{\iota\Omega\phi \star \left|D_i^+ + D_i^-\right|/2}{\iota\Omega\left|D_i^\pm\right|} \simeq \frac{\phi \star \left|D_i^\pm\right|/2}{\left|D_i^\pm\right|} \simeq \frac{\phi}{2} \ll 1. \quad (4.1)$$

In the linear case, $\phi \equiv \phi_L$ is independent of $\mathbf{D}$, so only the material parameters are constrained and the pulse properties play no role. In the nonlinear case, the criteria makes demands on the peak intensity of the pulse – but does not apply smoothness assumptions or bandwidth restrictions.

Secondly, linear and nonlinear vector terms from $\mathbf{U}$. These will have a criterion broadly the same as the scalar cases in eqn. (4.1), but with $\mathbf{U}$ replacing $\phi\mathbf{D}$. Thus for $i \in \{x, y, z\}$, we can write down constraints for each component U_i, which are

$$\iota\Omega\left|U_i\right|/2 \ll \iota\Omega\left|D_i^\pm\right| \quad \Longrightarrow \quad \left|U_i\right| \ll 2\left|D_i^\pm\right|. \quad (4.2)$$

In the linear case, $\mathbf{U} \equiv \mathbf{U}_L$, and since $\mathbf{U}_L$ and $\mathbf{D}$ have some linear relationship, this criterion only constrains the material parameters contained in $\mathbf{U}_L$, not the field profile or amplitude. The nonlinear case is similar to the scalar one, and the peak pulse intensity is restricted; e.g. for a $\chi^{(2)}$ medium, $|\mathbf{U}_N| \sim \chi^{(2)}|\mathbf{D}|$.

However, one complication of the vector cases is that a field consisting of only one field polarization component (e.g. D_x^+) may induce a driving in the orthogonal (and initially zero) components (e.g. $D_y^\pm$). Hence both $D_y^\pm$ fields will be driven

with the same strength, so that it is far from obvious that we can set D_y^- to zero, but still keep the D_y^+ without being inconsistent. However, as described above, it is the phase matching which ensures that forward residuals accumulate, whilst the non-matched backward residuals are subject to the RWA, and become negligible: hence we can still rely on eqn. (4.2), albeit under caution.

Thirdly, we have the residual terms indicated by the presence of time-varying currents. These are

$$\left|\dot{J}_i\right| \ll 2\Omega^2\left|D_i^\pm\right|. \quad (4.3)$$

Fourthly, we have the divergence terms $\mathbf{kk} \cdot \mathbf{D}$, $\mathbf{kk} \cdot \mathbf{P}_D$, which amount to charge density contributions either from $\mathbf{D}$ or $\mathbf{P}_D$, with $\sigma_P(k) = a_\varepsilon^{-1}\mathbf{k} \cdot \mathbf{P}_D$ and $\sigma_D(k) = \mathbf{k} \cdot \mathbf{D}$. Thus if the wavevector is oriented along a unit vector $\mathbf{u}$, then

$$\iota u_i\Omega\left|\sigma\right|/2k \ll \iota\Omega\left|D_i^\pm\right|$$
$$\sigma u_i \ll 2k\left|D_i^\pm\right|. \quad (4.4)$$

is the relevant criterion.

To summarize, these various criteria assert that modulations away from the reference evolution must be weak. Notably, weak nonlinearity is invariably guaranteed by material damage thresholds [31]. However, and in contrast to the usually favoured (but strictly speaking less causal) spatially propagated methods, the lack of any rigorous way to incorporate dispersion into the reference frequency $\Omega(k)$ makes it harder to satisfy the criteria for the ordinary linear response.

C. Uni-directional equation for $\mathbf{D}^+$

In the case where all of the slow-evolution criteria listed above hold, we can be sure that the backward directed field $\mathbf{D}^-$ is negligible, and if the no-accumulation condition also holds, then neither will there be any non-negligible backward evolving contributions to the field. Consequently, we can be sure that an initially negligible $\mathbf{D}^-$ remains so, and eqn.(3.2) becomes

$$\partial_t \mathbf{D}^+(\mathbf{k}) = +\iota\Omega(\mathbf{k})\mathbf{D}^+(\mathbf{k}) \quad + \frac{\iota\Omega(\mathbf{k})}{2a_\varepsilon}\mathbf{P}_D(\mathbf{D}^+, t; \mathbf{k})$$
$$- \frac{\dot{\mathbf{J}}(t; \mathbf{k})}{2\Omega(\mathbf{k})} \quad - \frac{\iota}{2\Omega(\mathbf{k})}\mathbf{k} \times \dot{\mathbf{M}}_B(\mathbf{D}^+, t; \mathbf{k})$$
$$- \frac{\iota\Omega(\mathbf{k})}{2}\frac{\mathbf{k}}{k^2}\mathbf{k} \cdot \mathbf{D}^+(\mathbf{k}) \quad + \frac{\iota\Omega(\mathbf{k})}{2a_\varepsilon}\frac{\mathbf{k}}{k^2}\mathbf{k} \cdot \mathbf{P}_D(\mathbf{D}^+, t; \mathbf{k}).$$
$$(4.1)$$

where now the polarization $\mathbf{P}_D$, diffraction, and divergence are solely dependent on the forward directed field $\mathbf{D}^+$.

D. Modifications

Just as for the comparable spatially propagated derivation [14] we might also apply some of the strategies used in other

approaches to get a more tractable and/or simpler evolution equation. Unlike in "traditional" derivations [32–35], none of these are required, but they nevertheless may be useful.

Briefly, a *co-moving frame* might be added, using $z' = z + t v_f$. This is a simple linear process that causes no extra complications; the leading RHS $\iota c D^+$ term is replaced by $\iota (c \mp v) D^\pm$. Note that setting $c = v$ will freeze the phase velocity of the pulse, not the group velocity. Also, a *carrier-envelope* separation could be implemented using $D(z) = A(z) \exp[\iota(\omega_1 t - k_1 z)] + A^*(z) \exp[-\iota(\omega_1 t - k_1 z)]$ defining the envelope $A(z)$ with respect to wave vector k_1 and carrier frequency ω_1. Multiple envelopes centred at different wave vectors and carrier frequencies (k_i, ω_i) could also be used to split the field up. This is typically done when the field principally comprises a number of distinct narrowband components (e.g. [29, 36, 37]). The wave equation can then be separated into one equation for each piece, coupled by the appropriate wavevector-matched polarization terms (cf [28]).

V. DISPERSION

Consider a dielectric slab waveguide of thickness d in the perpendicular x direction, propagation in the z direction, and with core and cladding permittivities ε_1 and ε_0. In the CW limit, the spatial structure of the waveguide gives rise to frequency dependent properties that are usually called "spatial dispersion". Even for this simple slab design, the dispersion relation has a non trivial form; in this case it is given by the solution to a transcendental equation [38]. Typically, this is written in a way implying we want to calculate k_x and $\beta = k_z$ from ω, i.e. for the transverse electric (TE) field modes we have

$$T(k_x d) = k_x^{-1} \sqrt{\omega^2 \mu_0 (\varepsilon_1 - \varepsilon_2) - k_x^2}, \qquad (5.1)$$

$$\text{where} \quad \beta^2(\omega) = \omega^2 \mu_0 \varepsilon_1 - k_x^2, \qquad (5.2)$$

and $T(k_x d)$ is either $\tan(k_x d)$ or $-\cot(k_x d)$.

However, we can now see that this is an implicitly spatially propagated representation, since in time propagated systems we do not (cannot) know ω, but we do know $\mathbf{k}$. Therefore eqn. (5.2) should be rewritten accordingly, using k_z for β (since wavevector is no longer a propagation reference), and Ω for ω (since frequency now *is* a propagation reference); thus

$$c^2 \Omega^2(k_x) = k_x^2 \frac{\varepsilon_1}{(\varepsilon_1 - \varepsilon_2)} \left[T^2(k_x d) + 1 \right]. \qquad (5.3)$$

$$\text{where} \quad k_z^2 = c^2 \Omega^2(k_x) - k_x^2. \qquad (5.4)$$

Here we see that the natural (reference) propagation frequency is defined solely by the transverse wavevector k_x, as is the matching evolution wavevector k_z. While apparently simpler than the traditional form, since Ω is directly given by k_x, we usually want to neglect any direct knowledge of the transverse properties. This means that what we really want is not $\Omega(k_x)$, but $\Omega(k_z)$, the inverse of the traditionally preferred $\beta(\omega)$.

Despite these complications, which are in any case rather similar to those we see in the traditional spatially propagated

picture, we can still find the same waveguide modes and their dispersions, but this time given as functions of k_z, not ω. So for some n-th mode of the waveguide, as expanded about some central evolution wavelength k_{zn}, which corresponds to a reference frequency $\Omega_n = \Omega(k_{zn})$, we might have

$$\Omega(k_z) = \Omega_n + \sum_m \gamma_m (k_z - k_{zn})^m, \qquad (5.5)$$

where $\gamma_1 = c + \bar\gamma_1$ has two contributions – that due to the vacuum (just c), and the estimated modifications due to spatial effects ($\bar\gamma_1$).

We may also wish to add the effects of whatever temporal response the material might have to our propagation calculation; but also decide that an approximate model where we modify the expression above in eqn. (5.5) is sufficent. This is the converse of the traditional procedure, where the waveguide dispersion is added to whatever $\beta(\omega)$ is appropriate for the medium's time response. Consequently we need only make the same assumptions: (a) that the two sources of dispersion are weak and so can simply be added, (b) that the temporal response is assumed to have negligible effect in the waveguide's propagating modes, and (c) that the temporal dispersion is the same at all points in the waveguide.

For simplicity we consider the usual situation where our dispersion has already been characterised and represented as a truncated power series in frequency. In line with the recommendations of [25], I exclude imaginary components (e.g. loss or gain terms) in either the wavevector reference β or the frequency reference Ω. Consequently, the reference wavevector considered here has the simple form

$$\beta(\omega) = \beta_0 + \kappa_1 (\omega - \omega_0) + \kappa_2 (\omega - \omega_0)^2, \qquad (5.6)$$

where $\kappa_1 = c^{-1} + \bar\kappa_1$ has two contributions – that due to the vacuum (just $1/c$), and the modification due to the temporal response ($\bar\kappa_1$).

I have chosen this form for $\beta(\omega)$ because it is possible to calculate its inverse, although in general – if there are more terms in the expansion, or β has some specific and non-trivial analytic form – the problem is more complicated. For this we again replace ω with Ω to denote the new role as a reference frequency, and replace β with k_z. With $\Omega_n = c k_{zn}$, this gives

$$\Omega(k_z) = \Omega_n - \frac{\kappa_1}{2\kappa_2} \left[1 \mp \sqrt{1 + 4(k_z - k_{zn}) \kappa_2 / \kappa_1^2} \right], \qquad (5.7)$$

which we could expand into a power series for small $k_{zn} - k_z$. Since κ_1 includes the vacuum, and dispersion κ_2 is assumed to be small, we should make the top sign choice since it gives the physically relevant small frequency corrections. The linear term in the series has the coeffient $\gamma_{1T} = 1/\kappa_1$ and the quadratic term $\gamma_{2T} = \kappa_2/\kappa_1^3$.

We now combine the two dispersive effects – the spatial/geometric/waveguide ones based on parameters γ_i, and converted temporal ones based on κ_i, to give a total reference

frequency based on k_z of

$$\Omega'(k_z) = \Omega_n + \bar{\gamma}_1 (k_z - k_{zn}) + \sum_{m>1} \gamma_m [(k_z - k_{zn})]^m,$$

$$- \frac{\kappa_1}{2\kappa_2} \left[1 - \sqrt{1 + 4(k_z - k_{zn}) \kappa_2 / \kappa_1^2} \right] \quad (5.8)$$

$$\simeq \Omega_n + \left(\bar{\gamma}_1 + \frac{1}{\kappa_1} \right) (k_z - k_{zn}) + \left(\gamma_2 - \frac{\kappa_2}{\kappa_1^3} \right) (k_z - k_{zn})^2. \quad (5.9)$$

Note that to avoid incorporating the vacuum contribution twice when summing the dispersive effects, I have replaced γ_1 with $\bar{\gamma}_1$.

In the complementary case, useful for spatial propagation, we can follow the converse – and indeed the *usual* procedure – to convert the CW effect of the waveguide's spatial structure $\Omega(k_z)$ into a frequency domain form [38, 39]. This then can be combined with temporal response of $\beta(\omega)$ to get

$$\beta'(\omega) = \beta_0 + \left(\bar{\kappa}_1 + \frac{1}{\gamma_1} \right) (\omega - \omega_0) + \left(\kappa_2 - \frac{\gamma_2}{\gamma_1^3} \right) (\omega - \omega_0)^2. \quad (5.10)$$

Discussion

We might have hoped to find a less approximate comparison between the spatially propagated and the temporally propagated approaches to dispersion what has just been discussed. Further, although qualitative comparisons can also be made, ideally we would like a specific and quantitative comparison as done for e.g. testing the unidirectional approximation [31].

That this is not as simple as it might sound can be demonstrated as follows: write two mathematically identical but physically and notationally distinct propagation equations, i.e.

$$\partial_z E(\omega) = +\iota K(\omega) E(\omega) \quad (5.11)$$

$$\partial_t D(z) = +\iota \Omega(z) D(z). \quad (5.12)$$

For simplicity, this comparison is done in the 1+1D limit and for unidirectional equations; the residual ω or z behaviour has been merged with the reference behaviour to give single propagation parameters $K(\omega)$ or $\Omega(z)$.

Superficially, the comparison between eqn. (5.11) and (5.12) looks promising. However, the vacuum part of $K(\omega)$ is ω/c, but the vaccum part of $\Omega(z)$ is in contrast a spatial derivative, i.e. $c\partial_z$. Even if it were possible to match up the functional forms of the non-vacuum temporal response of $K(\omega) - \omega/c$ with a spatial structure giving $\Omega(z) - c\partial_z$, the vacuum parts would remain unmatched. Alternatively, we could compare the z-propagated ω-domain to the t-propagated k-domain. However, although in this case the reference behaviours match, but the k-domain model now has a convolution that does not appear in the ω-domain one.

As a result, it is unclear how well approximating temporal dispersion as spatial properties works in practise; but then it is equally unclear how well approximating spatial properties as temporal dispersion works – even if this latter process is so ubiquitous in optics as to be effectively invisible. Removing either of the mismatches discussed above – either in reference behaviour or convolution – requires a narrowband (CW) limit, a process which necessarily obscures the parameter regime where interesting tests and comparisons can (and should) be made – i.e. the short pulse, large bandwidth limit. Indeed, the stringent nature of the CW approximation would defeat the entire point of the comparison.

Nevertheless, it may be possible to design numerical simulations containing structures and temporal responses whose scales are carefully graduated so that the mundane bandwidth issues can be kept separate from the interesting propagation ones. Such a design would, as desired, allow the propagation approximations (only) to be compared and contrasted by numerical simulation. However, this is a non-trival task, and one which I leave for later work.

VI. PULSE PROPAGATION

Third order nonlinearities are common in many materials, e.g. in the silica used to make optical fibres [18, 38, 39]. Here I compare 1+1D propagation such a waveguide, in order to elucidate the differences between the standard spatially propagated picture and the temporally propagated picture.

The assumptions made are those of transverse fields, weak dispersive corrections, and weakly nonlinear response; these all allow us to decouple the forward and backward wave equations. This decoupling then allows us, without any extra approximation, to reduce our description to one of forward only pulse propagation. The specific example chosen here is for an instantaneous cubic nonlinearity, but it is easily generalized to non-instantaneous cases or even other scalar nonlinearities.

A. Spatially propagated NLSE

The spatially propagated NLS for $E_x^+(\omega)$ based on a directional factorization [14] can be written

$$\partial_z E_x^+ = +\iota \beta E_x^+ + \sum_m \kappa_m \cdot (\omega - \omega_0)^m E_x^+$$
$$+ \frac{\iota k_0^2}{2\beta} \mathcal{F} \left[\chi^{(3)} E_x^2(t) E_x^+(t) \right], \quad (6.1)$$

where $\mathcal{F}[...]$ is the Fourier transform that converts the nonlinear polarization term into its frequency domain form. Here we have chosen a fixed, non-dispersive reference wavevector β, the temporal response of the propagation medium is encoded in the coefficients κ_j. For the temporal (time response) of the material, this is an in-principle exact re-representation of that response as a Taylor series expansion; although in practise the series is usually truncated after a few terms and the propagating fields restricted to within a bandwidth where only those terms are significant.

The important point to remember is that in this spatially propagated picture we do know the full time history, at least

in a computational sense – i.e. as we compute a solution to the propagation equation. This "computational past" [16], comprising all previously computed values of the field properties and material state means that a full frequency spectrum and response is known. As a result, the time response of the material *can* be directly encoded in an accurate and useful dispersion relation, and then be approximated as a Taylor series in ω.

Since this propagation is assumed to be taking place in a waveguide, we also have its geometric – wavelength sensitive – properties to consider. However, as discussed in the previous section we typically just re-represent those *as if they were instead generated by a temporal response* (see e.g. [38, 39]) Thus the coefficients κ_j usually combine both the effect of time response *and* geometric properties into a single total dispersion; this is the very basis of dispersion compensation in optical fibres, since the material (temporal) dispersion is offset by the waveguide design (spatial dispersion) [38].

B. Temporally propagated NLSE

Using the approach derived and presented in this paper, we can develop a time propagated counterpart to eqn. (6.1). In a purely mathematical sense, it is, apart from notation, identical to eqn. (6.1). However, in physical terms, the change is not trivial, since the physical meaning and boundary conditions differ significantly. Starting from eqn. (4.1) we can get a propagation equation in wavevector space for $D_x^+(k)$ which reads

$$\partial_t D_x^+ = +\iota\Omega D_x^+ + \sum_m \gamma_m \cdot (k_z - k_{z0})^m D_x^+ + \frac{\iota\Omega}{2a_\varepsilon} \mathcal{G}\left[\chi^{(3)} D_x^2(z) D_x^+(z)\right], \qquad (6.1)$$

where $\mathcal{G}[...]$ is the Fourier transform that converts the spatially localized nonlinear polarization into its wavevector domain form. *Here,* all dispersive behaviour has been combined into the coefficients γ_j, which mimic the common notion of spatial dispersion. Such terms can arise directly from a Taylor expansion in k of a relevant CW dispersion relation about some suitable reference wavevector; however, although we can use the expansion to represent the spatial structure of the propagation medium (e.g. a waveguide), the conversion from spatial structure to k expansion is only relatively straightforward in the CW limit.

If the propagation medium also has a temporal response (i.e. temporal dispersion), and if we do not want to model

it explicitly, then we can use the approach described in the previous section, where the temporal response of the material was (approximately) represented as if it were instead due to spatial properties.

Finally, note that in ordinary temporally propagated simulations, the computational past and the physical (temporal) past are identical; this is not the case for spatially propagated simulations, as mentioned above.

VII. CONCLUSION

I have derived a general first order wave equation for unidirectional pulse propagation in time that allows for arbitrary dielectric polarization, diffraction, and free electric charge and currents. This propagation equation has a utility all of its own, allowing efficient unidirectional propagation in a strictly causal manner (i.e. time directed). However, it requires comparable approximations to those used in deriving spatially propagated unidirectional propagation equations, although the specific tradeoffs are different.

The propagation equation was derived by first factorizing the second order wave equation for **D** into an exact bidirectional model, it applies the same type of well defined type of approximation to all non-trivial effects (e.g. nonlinearity, diffraction), and so reduces the propagation equations to a first order uni-directional wave equation. One nice feature of this factorization method is that it provides for a term-to-term comparison of the exact bi-directional theory with its approximate uni-directional counterpart.

In addition to its use where the causal simulation of a propagation problem is of strict concern, it also illuminates the rather shadowy process of handling material response and spatial structure in waveguides by constructing a single "combined dispersion" parameter. As discussed in section V, we can now see that this simple addition of temporal and spatial dispersions in an ad hoc manner stands on rather poor foundations, particularly in the few-cycle pulse limit. Further, it is hard to see any straightforward way of testing the limitations of this common procedure, which poses an interesting challenge for the future. That said, in most applications, we expect that a combined dispersion will remain a perfectly adequate approximation, since the pulses are rarely so short, and it seems reasonable to assume that the character of the pulse evolution should remain similar, even if details might differ. Lastly, the lessons learnt from this investigation could also be applied to other types of wave propagation, most notably acoustic systems (see e.g. [27]).

[1] T. Brabec and F. Krausz,
    Rev. Mod. Phys. **72**, 545 (2000),
    doi:10.1103/RevModPhys.72.545.
[2] P. B. Corkum,
    Phys. Rev. Lett. **71**, 1994 (1993),
    doi:10.1103/PhysRevLett.71.1994.
[3] K. J. Schafer, B. Yang, L. F. DiMauro, and K. C. Kulander,

    Phys. Rev. Lett. **70**, 1599 (1993),
    doi:10.1103/PhysRevLett.70.1599.
[4] T. Fuji, J. Rauschenberger, C. Gohle, A. Apolonski, T. Udem,
    V. S. Yakovlev, G. Tempea, T. W. Hansch, and F. Krausz,
    New J. Phys. **7**, 116 (2005),
    doi:10.1088/1367-2630/7/1/116.
[5] S. B. P. Radnor, L. E. Chipperfield, P. Kinsler, G. H. C. New,

Phys. Rev. A **77**, 033806 (2008),
doi:10.1103/PhysRevA.77.033806.

[6] J. Solanpää, J. A. Budagosky, N. I. Shvetsov-Shilovski, A. Castro, A. Rubio, and E. Räsänen,
Phys. Rev. A **90**, 053402 (2014),
doi:10.1103/PhysRevA.90.053402.

[7] R. R. Alfano and S. L. Shapiro,
Phys. Rev. Lett. **24**, 584 (1970),
doi:10.1103/PhysRevLett.24.584.

[8] R. R. Alfano and S. L. Shapiro,
Phys. Rev. Lett. **24**, 592 (1970),
doi:10.1103/PhysRevLett.24.592.

[9] J. M. Dudley, G. Genty, and S. Coen,
Rev. Mod. Phys. **78**, 1135 (2006),
doi:10.1103/RevModPhys.78.1135.

[10] D. R. Solli, C. Ropers, P. Koonath, and B. Jalali,
Nature **450**, 1054 (2007),
doi:10.1038/nature06402.

[11] A. Braun, G. Korn, X. Liu, D. Du, J. Squier, and G. Mourou,
Opt. Lett. **20**, 73 (1995),
doi:10.1364/OL.20.000073.

[12] S. L. Chin, S. Petit, F. Borne, and K. Miyazaki,
Jpn. J. Appl. Phys. **38**, L126 (1999).

[13] H. M. Milchberg, Y. H. Chen, Y. H. Cheng, N. Jhajj, J. P. Palastro, E. W. Rosenthal, S. Varma, J. K. Wahlstrand, S. Zahedpour,
Phys. Plasmas **21**, 100901 (2014),
doi:10.1063/1.4896722.

[14] P. Kinsler,
Phys. Rev. A **81**, 013819 (2010),
arXiv:0810.5689,
doi:10.1103/PhysRevA.81.013819.

[15] P. Kinsler (2012),
Directional pulse propagation in beam, rod, pipe, and disk geometries,
arXiv:1210.6794.

[16] P. Kinsler (2014),
Negative Frequency Waves? Or: What I talk about when I talk about propagation,
arXiv:1408.0128.

[17] P. Kinsler,
Eur. J. Phys. **32**, 1687 (2011),
arXiv:1106.1792,
doi:10.1088/0143-0807/32/6/022.

[18] G. P. Agrawal,
Nonlinear Fiber Optics
(Academic Press, Boston, 2007), 4th ed.,
ISBN 978-0-12-369516-1.

[19] L. Berge and S. Skupin,
Discrete Continuous Dyn. Syst. (DCDS-A) **23**, 1099 (2009),
doi:10.3934/dcds.2009.23.1099.

[20] K. S. Yee,
IEEE Trans. Antennas Propagat. **14**, 302 (1966), fDTD,
doi:10.1109/TAP.1966.1138693.

[21] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson,
Comput. Phys. Commun. **181**, 687 (2010),
doi:10.1016/j.cpc.2009.11.008.

[22] S. J. Carter,
Phys. Rev. A **51**, 3274 (1995),
doi:10.1103/PhysRevA.51.3274.

[23] A. Ferrando, M. Zacares, P. F. de Cordoba, D. Binosi, and A. Montero,
Phys. Rev. E **71**, 016601 (2005),
doi:10.1103/PhysRevE.71.016601.

[24] K. J. Blow and D. Wood,
IEEE J. Quantum Electronics **25**, 2665 (1989),
doi:10.1109/3.40655.

[25] P. Kinsler,
Phys. Rev. A **79**, 023839 (2009),
arXiv:0901.2466,
doi:10.1103/PhysRevA.79.023839.

[26] P. Kinsler, S. B. P. Radnor, and G. H. C. New,
Phys. Rev. A **72**, 063807 (2005),
arXiv:physics/0611215v1,
doi:10.1103/PhysRevA.72.063807.

[27] P. Kinsler (2012),
Acoustic waves: should they be propagated forward in time, or forward in space?
arXiv:1202.0714.

[28] P. Kinsler, G. H. C. New, and J. C. A. Tyrrell (2006),
Phase sensitivity of perturbative nonlinear interactions,
arXiv:physics/0611213.

[29] R. W. Boyd,
Nonlinear Optics
(Academic Press, New York, 2008), 3rd ed.,
ISBN 978-0-12-369470-6.

[30] S. E. Harris,
Appl. Phys. Lett. **9**, 114 (1966),
doi:10.1063/1.1754668.

[31] P. Kinsler,
J. Opt. Soc. Am. B **24**, 2363 (2007),
arXiv:0707.0986,
doi:10.1364/JOSAB.24.002363.

[32] T. Brabec and F. Krausz,
Phys. Rev. Lett. **78**, 3282 (1997),
doi:10.1103/PhysRevLett.78.3282.

[33] P. Kinsler and G. H. C. New,
Phys. Rev. A **67**, 023813 (2003),
arXiv:physics/0212016v1,
doi:10.1103/PhysRevA.67.023813.

[34] M. Geissler, G. Tempea, A. Scrinzi, M. Schnürer, F. Krausz, and T. Brabec,
Phys. Rev. Lett. **83**, 2930 (1999),
doi:10.1103/PhysRevLett.83.2930.

[35] M. Scalora and M. E. Crenshaw,
Opt. Comm. **108**, 191 (1994),
doi:10.1016/0030-4018(94)90647-5.

[36] P. Kinsler and G. H. C. New,
Phys. Rev. A **69**, 013805 (2004),
arXiv:physics/0411120,
doi:10.1103/PhysRevA.69.013805.

[37] P. Kinsler and G. H. C. New,
Phys. Rev. A **72**, 033804 (2005),
arXiv:physics/0606111,
doi:10.1103/PhysRevA.72.033804;
see also arXiv:physics/0606112.

[38] K. S. Thyagarajan and A. Ghatak,
Fiber Optic Essentials
(Wiley-IEEE Press, 2007),
ISBN 978-0-470-09742-7,
doi:10.1002/9780470152560.

[39] J. K. Shaw,
Mathematical Principles of Optical Fiber Communications,
CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM, 2004),
ISBN 978-0-89871-708-2,
doi:10.1137/1.9780898717082.

Appendix A: Factorizing

Here I present a simple overview of the mathematics of the factorization procedure, since full details can be found in [23]. In the calculations below, I alter the physics but not the mathematical process (of [23]) to instead transform into frequency space, where the t-derivative ∂_t is converted to $\iota\omega$. Also, we have that $\Omega^2 = c^2 k^2 / n^2$, and the unspecified residual term is denoted Q. The second order wave equation can then be written

$$\left[\partial_t^2 + \Omega^2\right] D = -Q \qquad (A.0.1)$$

$$\left[-\omega^2 + \Omega^2\right] D = -Q \qquad (A.0.2)$$

$$D = \frac{1}{\omega^2 - \Omega^2} Q \qquad = \frac{1}{(\omega - \Omega)(\omega + \Omega)} \qquad (A.0.3)$$

$$= \frac{-1}{2\Omega}\left[\frac{1}{\omega + \Omega} - \frac{1}{\omega - \Omega}\right] Q. \qquad (A.0.4)$$

Now $(\omega - \Omega)^{-1}$ is a forward-like (Green's function) propagator for the field, but note that in my terminology, it *evolves* the field. The complementary backward-like propagator is $(\omega + \Omega)^{-1}$. As already described in the main text, we now write $D = D^+ + D^-$, and split the two sides up to get

$$D^+ + D^- = \frac{-1}{2\beta}\left[\frac{1}{\omega + \Omega} - \frac{1}{\omega - \Omega}\right] Q \qquad (A.0.5)$$

$$D^{\pm} = \frac{\pm 1}{2\Omega}\frac{1}{\omega \mp \Omega} Q \qquad (A.0.6)$$

$$\left[\omega \mp \Omega\right] D^{\pm} = \pm \frac{1}{2\Omega}\frac{1}{\omega \mp \Omega} Q \qquad (A.0.7)$$

$$\iota\omega D^{\pm} = \pm \iota\Omega D^{\pm} \pm \frac{\iota}{2\Omega} Q. \qquad (A.0.8)$$

Finally, we transform the frequency space $\iota\omega$ terms back into normal time to give t derivatives, resulting in the final form

$$\partial_t D^{\pm} = \pm \iota\Omega D^{\pm} \pm \frac{\iota}{2\Omega} Q. \qquad (A.0.9)$$

Appendix B: The no accumulation approximation

In the main text, I describe the no accumulation approximation in spectral terms as a RWA approximation. However, it is hard to set a clear, accurate criterion for the RWA approximation to be satisfied in the general case, since it requires knowledge of the entire propagation before it can be justified. In this appendix, I take a different approach to determine the conditions under which the approximation will be satisfied.

First, consider a forward evolving field so $D = D_0 \exp(\iota\omega t)$, and therefore

$$D_0^- = \frac{\omega - \Omega}{\omega + \Omega} D_0^+ \qquad = \xi D_0^+, \qquad (B.0.1)$$

where as noted ω can be difficult to determine, and may even change dynamically; here we can assume it corresponds to the propagation wave vector that would be seen at if all the conditions holding at a chosen position also held everywhere else. On this basis, we can (might) even (try to) define $\omega = \omega(t)$, where by analogy to the linear case we might assert that $\omega^2(t) = \Omega^2 + Q(t)/D(t)$, so that for small Q, we have $\omega D \simeq \Omega(D + Q/2\Omega^2)$.

Let us start by assuming our field is propagating and evolving forwards (only), with perfectly matched $D^{\pm}$ fields; so that $D^- = \xi D^+$. but then it happens that Q changes by δQ over a small interval δt, likewise ξ changes by $\delta\xi$. The $D^{\pm}$ will no longer be matched, and now the total field splits into two parts that evolve in opposite directions. The part that continues to evolve forward has D^+ nearly unchanged, but the forward evolving D^- has changed size (and is now $\propto (\xi - \delta\xi)$) to stay perfectly matched according to the new Q. The rest of the old D^- ($\propto \delta\xi$) now propagates backwards, taking with it a tiny fraction of the original D^+ (and is $\propto \xi\delta\xi$).

Comparing the two backward evolving D^- components at t and $t + \delta t$, and taking the limit $\delta t \to 0$ enables us to estimate that the backward evolving D^- field changes according to

$$\partial_t D_{0,backward}^- = \frac{2\Omega}{(\omega + \Omega)^2}\left[\partial_t \omega\right] D_{0,forward}^+. \qquad (B.0.2)$$

Now, using the small-Q approximation for ω, we can write

$$\partial_z D_{0,backward}^- = \frac{1}{(\omega + \Omega)^2}\left[\partial_t Q\right] e^{-\iota\omega t}. \qquad (B.0.3)$$

where the exponential part removes any oscillations due to the linear part of Q; i.e. if $Q = \chi D$ then

$$\partial_t D_{0,backward}^- = \frac{1}{(\omega + \Omega)^2}\left[\partial_t \chi\right]. \qquad (B.0.4)$$

So here we see that backward evolving fields are only generated from forward evolving fields due to changes in the underlying conditions (i.e. either material response or pulse properties), but that for the reflection to be strong those changes will have to be significant on the order of a period, or be periodic so that phase matching of the the backward wave could occur.